

Contents

1	New Trends in Web User Behaviour Analysis	1
	<i>Pablo E. Román, Juan D. Velásquez, Vasile Palade, Lakhmi C. Jain</i>	
1.1	Introduction	1
1.2	The Web Operation	2
1.3	The Big Challenge: To Understand the Human Behaviour by Data Analysis	4
1.4	Novel Approaches for Modeling the Web User Behaviour	6
1.5	Characterizing the Web User Browsing Behaviour	7
1.6	Real-World Applications	8
1.7	Chapters Included in the Book	8
	References	9
2	Web Usage Data Pre-processing	11
	<i>Gaston L'Huillier, Juan D. Velásquez</i>	
2.1	Introduction	11
2.2	Web Usage Data Sources and Collection	12
2.2.1	Nature of Web Usage Data	13
2.2.2	Merging and Cleaning Data	15
2.3	Web Session Reconstruction and User Identification	15
2.3.1	Heuristics and Methods for Sessionization	16
2.3.2	Dealing with Incomplete Information	21
2.3.3	Model Assessment for Sessionization Algorithms	22
2.4	Profiling and Relevance Data Processing	23
2.4.1	Relevance Feedback and Explicit Ratings	24
2.4.2	Eye Tracking Data Processing	26
2.5	A Discussion on Privacy in the World Wide Web	29
2.6	Summary	30
	References	31

3	Cognitive Science for Web Usage Analysis	35
	<i>Pablo E. Román, Juan D. Velásquez</i>	
3.1	Introduction	35
3.2	Theories of Preferential Decision Making	37
3.2.1	Tversky's Elimination by Aspect (EBA)	38
3.2.2	Prospect Theory	38
3.2.3	Expected Utility Theory (EUT)	40
3.2.4	The Random Utility Model	40
3.2.5	Multiple Resource Theory (MRT)	42
3.2.6	Scent-Based Navigation and Information Foraging in the ACT Architecture (SNIF-ACT) Model	43
3.2.7	CoLiDeS Model	44
3.2.8	Other Models	44
3.3	Neurocomputing Theories	45
3.3.1	Neuroscience Foundations	46
3.3.2	Biological Insight on Decision Making	47
3.3.3	Wiener Diffusion Process	51
3.3.4	Decision Field Theory (DFT) and the Ornstein-Uhlenbeck Process	54
3.3.5	Leaky Competing Accumulator Model (LCA)	55
3.3.6	Other Models	56
3.4	A Neurocomputing Application to the Web Usage Analysis	57
3.4.1	The Time Course of the Web User	58
3.4.2	A Model From Psychology	60
3.4.3	A Random Utility Model for Text Preference	63
3.4.4	Differential Equation for Probability Distributions	64
3.5	Discussion	67
	References	68
4	Web Usage Mining: Discovering Usage Patterns for Web Applications	75
	<i>Giovanna Castellano, Anna M. Fanelli, Maria A. Torsello</i>	
4.1	Introduction	75
4.2	The WUM Process	77
4.2.1	Web Usage Data Collection	78
4.2.2	Web Usage Data Preprocessing	80
4.2.3	Web Usage Pattern Discovery	84
4.2.4	Application of Web Usage Patterns	89
4.3	A Fuzzy Relational Clustering Technique for Mining User Categories	91
4.3.1	Computing (dis)similarity among Web Users	92
4.3.2	Grouping Users by Fuzzy Clustering	94
4.3.3	A Case Study	95
4.4	Conclusions	99
	References	99

5	Web Opinion Mining and Sentimental Analysis	105
	<i>Edison Marrese Taylor, Cristián Rodríguez O., Juan D. Velásquez, Goldina Ghosh, Soumya Banerjee</i>	
5.1	What Is Web Opinion Mining (WOM)?	105
5.2	How to Do WOM?	107
5.2.1	Aspect-Based Opinion Mining	107
5.2.2	Document Level Opinion Mining	109
5.3	Sources for Web Opinion Mining	109
5.3.1	Blogs, News and Forums	110
5.3.2	Twitter	111
5.3.3	Other Media	112
5.4	Visualization Techniques	112
5.5	An Application of WOM in Twitter	114
5.5.1	Extracting Significant Information from Tweets	115
5.5.2	Data Quality Evaluation	116
5.5.3	Analysis and Results	119
5.5.4	Tweeter as Novel Web User Behavior Data Source	123
5.6	Summary	123
	References	124
6	Web Usage Based Adaptive Systems	127
	<i>Pablo Loyola Heufemann, Jorge Gaete Villegas, In-Young Ko</i>	
6.1	Introduction	127
6.1.1	An Overview on Web Usage Based Adaptive Systems	128
6.1.2	Web Usage Based Adaptive Systems Evolution	129
6.2	Process Framework	129
6.2.1	Domain Model	131
6.2.2	User Model	132
6.2.3	Adaptation Model	133
6.3	Existing Approaches for Web Usage Based Systems	133
6.3.1	Content Adaptation	133
6.3.2	Navigation Adaptation	136
6.3.3	Presentation Adaptation	140
6.4	Privacy Issues in User Data Collection and Usage	141
6.4.1	Generation, Access and Usage of Data	142
6.4.2	What Is Private Data?	143
6.4.3	Privacy-Friendly Development Policy	144
6.5	Future Challenges	144
	References	146
7	Recommender Systems: Sources of Knowledge and Evaluation Metrics	149
	<i>Denis Parra, Shaghayegh Sahebi</i>	
7.1	Introduction	149
7.2	Classification of Recommender Systems	151
7.2.1	Rule-Based Recommender Systems	152

7.2.2	Content-Based Recommender Systems	153
7.2.3	Collaborative Filtering Recommender Systems	154
7.2.4	Hybrid Recommender Systems	155
7.3	Sources of Knowledge for Recommender Systems	156
7.3.1	Ratings	156
7.3.2	Implicit Feedback	156
7.3.3	Social Tags	157
7.3.4	Online Social Networks	158
7.3.5	Context	159
7.3.6	Heterogeneous Sources of Knowledge	160
7.4	Evaluation Metrics for Recommender Systems	160
7.4.1	Prediction-Based Metrics	161
7.4.2	Information Retrieval Related Metrics	161
7.4.3	Diversity, Novelty and Coverage	163
7.4.4	Implicit Feedback and Partial Knowledge of User Preferences	164
7.4.5	Beyond Metrics: User Study Frameworks	165
7.5	Web Recommendations	166
7.5.1	Sources of Knowledge for Web Recommendation	166
7.5.2	Methods for Web Recommendation	167
7.5.3	Evaluation Metrics for Web Recommendation	167
7.6	Summary	168
	References	168
	Index	177